

Comments on: Dynamic priority allocation via restless bandit marginal productivity indices

I.J.B.F. Adan · O.J. Boxma

Published online: 27 September 2007
© Sociedad de Estadística e Investigación Operativa 2007

There is an abundance of real-life situations of the following nature. Several projects, of various types, need to be done by some person, machine, or other device. These require certain amounts of time, while costs (or rewards) are associated with performing such projects, delaying them, or switching between projects. The problem now is: Find a policy that determines, for given numbers of projects of the various types, on which project one should work, so as to minimize (discounted or average) costs. Examples of this problem turn up in administrative settings, in manufacturing, in computer-communications, etc.

The problem of dynamically prioritizing projects (traffic classes) may often be formulated as a Markov decision problem, but the state space explosion makes the numerical solution of realistically sized problems usually prohibitive. An important class of heuristic policies, which often provide near-optimal solutions (and sometimes even can be shown to be optimal), is the class of priority index policies. A static index rule employed in minimizing completion costs in single-machine scheduling is the ratio of holding cost rate to processing time: work on the job with highest index. A dynamic index rule has been shown (Gittins and Jones (1974), as referenced in Niño-Mora 2007) to be optimal for the multi-armed bandit problem, viz., sequential allocation of work to a collection of stochastic projects (=bandits), so as to maximize the expected total discounted reward earned over an infinite horizon. Whittle (1988) discusses the multi-armed *restless* bandit problem, i.e., bandits can change state while being passive, under the long-run average criterion.

This comment refers to the invited paper available at: <http://dx.doi.org/10.1007/s11750-007-0025-0>.

I.J.B.F. Adan (✉) · O.J. Boxma

EURANDOM and Department of Mathematics & Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: iadan@win.tue.nl

The index proposed by Whittle is suboptimal, and it is asymptotically optimal under certain conditions. However, existence of this index is not always guaranteed; as Niño-Mora points out, indexable projects are precisely those that obey the classic economic law of diminishing marginal returns, or diminishing marginal productivity, in such a way that there is a well-defined marginal value of work at every state.

In a series of papers, Jose Niño-Mora has introduced and explored the concept of *marginal productivity index* (MPI). It provides a unifying framework for the indices mentioned above. The MPI values have an appealing geometric interpretation: MPI measures the marginal values of work or slopes in the upper boundary of the achievable performance region. This interpretation provides an economic justification for using MPI in a multi-project scenario: MPI policies aim to “dynamically allocate work to those projects that can make better use of it”, as Niño-Mora puts it.

In the paper of Niño-Mora (2007), Niño-Mora surveys his own recent work on the use of MPI to dynamic prioritization for restless bandit problems. He reviews theoretical developments and algorithmic aspects, and discusses several applications to admission control, routing to parallel queues, and production scheduling.

MPI appears to be a very fruitful and powerful concept. With a wide range of examples, Niño-Mora demonstrates the applicability of his ideas to relevant real-life problems; applications can be found in manufacturing (Sect. 3.2) and wireless communication (Sects. 4.3, 4.4), but also in, e.g., machine maintenance (Glazebrook et al. 2005) and outsourcing of warranty repairs (Opp et al. 2005). Numerical studies like his (2007b) paper show that his MPI policy is often near-optimal and considerably better than existing policies. In several publications from 2007, he develops MPI extensions, tackling such issues as (i) control of systems with delayed information, (ii) multi-armed bandits with switching costs, (iii) with switching delays, and (iv) with deadlines—hence, with a finite planning horizon. It is worthwhile to mention that extensions such as (ii) and (iii) are relevant in multi-product production systems, where switching from one product type to another requires a (large) set-up time (or cost), and in dynamically controlled traffic light systems, where switching from one traffic stream to another requires a clearance time.

The survey (Niño-Mora 2007) is the accumulation of more than a decade of deep and systematic research, mainly performed by Niño-Mora on his own. Incidentally, although the primary goal of Niño-Mora (2007) was to survey recent work of the author himself, it would have been interesting and useful to position the work of the author in relation to recent work of other researchers in this field. He has carefully set out a very ambitious research program (to which he has referred in several lectures). Even more remarkably, he has stuck to that program, meticulously filling the various gaps needed to complete his building. In performing this program, he has often been led by key concepts and ideas by pioneers in the field. With a good sense of historical perspective, he mentions in this respect the classical ideas of Smith (1956) on scheduling, by Bradt et al. (1956) and Bellman (1956) on the one-armed bandit, by Gittins and Jones (1974) on the Gittins index, Klimov’s (1974) idea to compute an index policy via an algorithm based on LP duality, Whittle’s (1988) index approach to the restless bandit problem, and even the application of the marginalist ideas from economics by Koopman (1957) and Kantorovich (1959) to resource allocation theory.

Completing this program, and writing such a clear and comprehensive survey about it, is a remarkable achievement. The wealth of ideas which are outlined could

also have given rise to several PhD projects; in that respect, it is a pity that the project has not yet led to much more training of students and collaboration with other senior researchers, which could have moved this important work ahead at a higher speed. As the author remarks at the end of his survey, many interesting issues remain to be explored. Issues that might be added are (i) a systematic comparison with other approaches to develop effective, simply structured index policies such as dynamic programming one-step policy improvement (see, e.g., Opp et al. 2005; Sassen et al. 1997) and (ii) a comparison with the ‘more traditional’ static or dynamic optimization of polling systems (see e.g., Borst 1994; Weststrate 1992; Yechiali 1991).

References

- Borst SC (1994) Polling systems. PhD thesis, University of Tilburg, Tilburg
- Glazebrook KD, Mitchell HM, Ansell PS (2005) Index policies for the maintenance of a collection of machines by a set of repairmen. *Eur J Oper Res* 165:267–284
- Opp M, Glazebrook KD, Kulkarni VG (2005) Outsourcing warranty repairs: dynamic allocation. *Nav Res Logist* 52:381–398
- Niño-Mora J (2007) Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, this issue
- Sassen SAE, Tijms HC, Nobel RD (1997) A heuristic rule for routing customers to parallel servers. *Stat Neerl* 51:107–121
- Weststrate JA (1992) Analysis and optimization of polling models. PhD thesis, University of Tilburg, Tilburg
- Yechiali U (1991) Optimal dynamic control of polling systems. In: Cohen JW, Pack CD (eds) *Queueing, performance and control in ATM*. North-Holland, Amsterdam, pp 205–217